Submitted by
Soumitra Sengupta, PhD
Information Security Officer, NewYork-Presbyterian Hospital, New York
Information Security Officer, Columbia University Medical Center, New York
Assistant Clinical Professor, Dept of Biomedical Informatics, Columbia University, New York

HIT Standards Committee
Hearing on Health Information Technology Security Issues, Challenges, Threats, and Solutions

Data Theft, Loss, and Misuse Panel – to address security challenges involving accidental loss of data, data theft, extortion and sabotage, including criminal activities and other related areas.

Questions:

1. **Briefly describe your organization and your information security approach to data theft, loss, and misuse.**

   NewYork-Presbyterian (NYP) Hospital is a tertiary care, academic medical center in New York City. It is intimately affiliated with two medical Schools: Cornell University (Weill Medical College) and Columbia University (College of Physicians & Surgeons, Nursing School, College of Dental Medicine). The Hospital has 5 large inpatient centers in Manhattan and Westchester, and is affiliated with several acute care, long-term care, and specialty care institutions in New York tri-state geographic area. In 2008, the Hospital had 2242 beds with about 120,000 discharges, and over 1.6 million outpatient and emergency visits.

   There are over 30,000 user accounts and around 75 major applications in the Hospital. The computer network itself has about 10,000 nodes (5 Class B IP networks and a private Class A IP network), with 2 high speed connections to Internet 2 through the Universities. There are 5000 attending physicians (and equivalent), 2000 residents and fellows, and 2000 students. The main clinical repository holds more than 2.5 million patient records.

   Our security posture is based on several administrative and technical controls. We require annual training and distribute periodic reminders to the users. Usual controls such as firewalls, Intrusion detection/prevention systems, network forensics, VPN, etc. are implemented. To address data loss, we are in the midst of encrypting all institutional laptops, requiring purchase of encrypted USB drives, and ensuring encryption of all tape backup. Data theft and misuse rely significantly on audit log controls and investigations. A new initiative on Data Leakage Prevention is currently at the vendor selection stage, and a Role and Identity management system will be implemented next year.

2. **Provide one or two examples of information security issues you have faced recently related to data theft, loss, and misuse, and describe how you addressed these issues.**

   There are several examples of data theft and misuse that we have observed in our environment. A large scale incident in February 2008 related to an identity theft operation. We are concerned about not infrequent external Internet based Bot activities which have a potential for large scale problem.

   Otherwise, the common misuse of small scale clinical data occurs for celebrity patients (e.g., baseball players, media celebrities, Latin artists, ex-President) and for employees or relatives of employees. Increasingly, we notice theft of clinical data in domestic disputes and child custody cases. We are also concerned about records of New York industry leaders who may not be celebrities, but a breach of their records may have a large impact.

   In Feb 2008, upon queries from the FBI, we investigated a patient registrar who was found to be screen-printing demographic information of patients in our Registration system and was handing over the paper stack for small sums of money. Although we had audit log collection and rudimentary alerting based on volume per day in place, our alert mechanism failed due to mistakes in log volume calculations and unrefined false positives. The logs were consulted for access by this person, and in an expansive count including all years of activities, we informed over 48,000 patients that their identity data may have been compromised, and offered credit protection service for 2 years. Subsequently, NYP senior management met with the District Attorney's office with explanation of inherent richness of healthcare demographic data, need for access by employees such as registrars, and difficulty of need-to-know and minimum necessary.

   Subsequently, our audit log alerting is now much improved. We currently trigger alerts on specific conditions such as (1) Number of consecutive medical record numbers accessed by a user, (2) Sudden significant change in number of records accessed by a user compared to their own past practice, (3) a significant variation in number of records accessed per day within a group of users with the same job title, (4) Number of hours an account is in use in a day, etc. Such alerts are investigated with the management of the user and reasons and explanations are documented. The metrics of the process are reported to the Audit board of NYP. Currently, we have 30 applications reporting about 700,000 log records (130MB raw data) for about 65,000 patients each day in our audit log server.

   We should also note that our log volume miscalculations are direct result of confusing, cryptic and non-standard audit logs of the registration system which uses medical record numbers sometimes and uses an internal identification number at other times. The logs also do not include update activities.

We initiated an NYP-wide Information Security Enhancement project under the senior management leadership, and engaged a security consultant to improve our security posture. The exercise has resulted in 18 major tasks, which include coordination among institutions, reduction of Personally Identifiable Information (PII), additional policies and controls, and emphasis on metrics based on ISO 27002. The project status is reported to the Information technology and Audit boards of the NYP.

3. **What kinds of trade-off's have you had to make between security and usability, and other operational considerations?**

We require one person one userid in our applications, and strongly resist generic userids. Nevertheless, all systems have internal generic ids that we require to be documented. Over time, we have negotiated with vendors (this is a slow, costly, and resource intensive activity, and therefore not a 100% compliant) to change their authentication to our central directory server, which has resulted in decent reduced sign-on environment. A direct benefit is that password strength (we require complex passwords) and password expiry/change activities (every 6 months) have less usability cost for our users. While we are aware of a 10-20% of users seeking service desk help for password changes, we are reassured that majority has learnt to use a password synchronization system to change and select strong passwords. The institution has thus spent significant resources to address usability and security up front.

We should note that vendors are not eager to consider themselves as partners and their systems as co-residents in our computing environment in order to reduce our overall operational costs, improve usability for sign ons, and improve security at the same time. They are not sure which specific authentications they should offer and support, and end up doing the least, unless this is negotiated as part of purchase. Another complicating factor is the outsourced applications which mostly do not offer federated authentications, thus negating the outsourcing advantage with extra usability and account management costs. A very clear set of standards are needed to address this issue.

The other usability issues are that of application timeouts and monitor placements in locations such as Emergency department. The former varies from 5 minutes through 30 minutes in different work areas. The latter requires new technologies such as proximity cards, which we have not deployed, partly due to lack of standards that will work for more than one application or one vendor solution.

We follow the principle that care trumps security, and in each case we consider whether care is impacted in a large, generic way, and if so, we make a collective determination towards the acceptable level of risk that the institution is willing to take.

4. **What information security standards are you currently using to protect your business from data theft, loss, and misuse?**

    1. Education and reminders. In addition to periodic changes to the content, we will implement 'hands-on' training. For example, we will use a solution that sends 'Phishing email' to users, and if they react with lack of understanding, we will respond back to them with immediate corrective education.
    2. Due to NY State breach notification laws, we implemented solutions to scan for PII exchanges onto the Internet using a network forensics tool. Our Security Enhancement project also identified the clear need for a robust Data Leakage Prevention solution which is being implemented and will be instrumental towards PHI protection as well. The standards in this area will be developed over time – which users are allowed to copy PII or PHI to removable media and why, to what degree email may be used to exchange information, etc.
    3. We convert audit logs from various systems into a canonical format with following information –
        a. Date and time
        b. User identification (User id/Employee Id)
        c. Workstation/Client identification (IP address/NetBIOS name/etc.)
        d. Patient identification (MRN/EMPI number/Internal number)
        e. Action description (read/wrote/print/update/etc.)
        f. Data description (lab/radiology/pathology/orders/demographics/etc)

        This allows the Security Office to create a proper alerting mechanism, and develop a consistent audit log review application. The Patient Relations department uses the system to observe accesses for celebrity patients. The Privacy Office uses the system to investigate a patient complaint.
    4. We receive weekly feeds from several Threat Intelligence groups (Shadow servers, Emerging threats, SURBL, etc.) to identify malicious IP addresses as known Command & Control Bot servers and we feed these addresses and subnets into our firewalls to prohibit access – over 5000 IP addresses and subnets are thus blocked. These groups create de facto standard lists of malware sites.

5. **What challenges have you had to address in implementing these standards (e.g., training)?**

    There are no security standards that are black and white, and we are severely dependent on what vendors offer and support. Vendors come in all sizes – device manufacturers are excited about offering wireless devices, and we find their security ends up being mostly Pre-Shared Keys, which are not very secure in the long run. Healthcare institutions do not necessarily have a great culture of technical prowess in matters such as Public Key Infrastructure. Lack of expertise in terms of knowledgeable people is a significant barrier to information security, especially when we compare it with shortage of nursing, or other issues with healthcare in general. Lack of resources is also a

significant barrier because good security requires good controls, some of which are complex and expensive.

Education issues ultimately relate to how much detail that can be presented before it becomes too confusing. We would like to reduce SSN for our patients except that Medicare numbers reflect the SSN by default. This one issue alone would reduce the attractiveness of health care demographics to identity thieves. We are learning to be careful with credit cards due to PCIDSS. Lack of complete and understandable audit logs will hamper the efforts of understanding what is appropriate for access based on past data will continue for some time. Vendors need to be more accountable to offer automated provisioning and de-provisioning, and automated role assignment externally so that we can centrally manage accesses with accountability; the current, mostly manual account management leads to errors, delays, and lack of segregation of duty and audits.

6. **What is the role/value of interoperable information security standards in helping to protect your business from data theft, loss, and misuse?**

Healthcare by definition is a heterogeneous collection of applications, as reflected by the practice of care itself with its specialties and subspecialties – the information system environment simply follows the practice of care. We find systems in Operating Rooms, ICUs, specialty wards such as Transplant, Inpatient, Outpatient, Scheduling, Billing, Referrals, specific Radiology/ Cardiology/OB/Gyn/etc. devices and their alerting and analysis systems, myriad of wireless Point-of-Service devices and their backend servers, Medication distribution systems, Patient access to records through PHR, and Health Information Exchanges, distribution of data to City, State and Federal agencies.

With confidentiality and integrity being the first two components of information security, if we focus on controls related to authentication, authorization, audit logs, and encryption across these diverse care solutions, each critical to a group of clinical users, to address data theft, loss or misuse, we require clear set of standards to implement security that can be measured and compared across institutions. Then by correlating actual and known data theft, loss, and misuse cases with existing security metrics, we can assign likelihoods of problems, which can then help us acquire the desired change in controls, and/or correct risk mitigation and transference methods to manage the environment.

7. **What are the current limitations or gaps in interoperable information security standards addressing data theft, loss, and misuse?**

   One primary issue is that vendors so far do not have an incentive to follow standards because other than HL7 and DICOM, there are not many health care standards defined as yet. Also, information security is more than healthcare issue, and again there are no well-defined standards that an institution can easily follow. Consider NIST set of standards, which are a very large set of requirements, and the complexity of implementing each control is significantly technical and administrative, which unfortunately works against the collective knowledge and purpose of IT organizations in healthcare institutions which are looking to enable more technological solutions to help care and business. Not having clear procedures and standards to follow permits variable levels of controls, resulting in lesser security. A survey of how systems are acquired in healthcare institutions may be illuminative in demonstrating lack of best practice standards to evaluate value, operational efficiency, and security of acquired systems. To what degree configuration of servers, workstations, and devices conform to a tight standard in large academic medical centers is also indicative of limitations of decentralized information technology management within institutions. There is a fundamental mismatch between healthcare and say banking industry in how the data are shared and valued, and the basic tenet for the data to be used for care, education and research. We require standards that work for healthcare which are not the same as the banking security standards (if they had one). Process standards are more important than technological standards: what access rights should students have, how can researchers get access to data for research, what level of de-identification have to be implemented, but with ability to re-identify, for the research or quality data, and so on.

8. **What new and emerging issues around data theft, loss, and misuse do you foresee over the next 2-3 years?**

   The impact of a public web site that lists PHI losses may show that this is larger problem than we think it is. The device proliferation in the wireless world with weak security may cause significant data mischief by external agents. At an extreme, ARRA/HITECH breach notification may act as an incentive for a Bot master to blackmail healthcare institutions by holding the data hostage (say by encrypting) even within our own servers. We are concerned about how to reduce employee accessing other employee record – although accounting of disclosure in ARRA/HITECH may be the answer. The flipside problem of the same is how to explain why hundreds of users access clinical records for a couple of days of inpatient stay, and that they are not stealing or misusing the data. Perhaps focus should shift to severe penalties on people who abuse the data for personal gain, as opposed to penalties on institutions or individuals for accidentally losing the data.